

Section 1: Collecting data

Exercise level 1

1. Name the sampling method being used in each of the following cases. Choose the sampling method from the list below.

Opportunity sample	Stratified sample	Self-selected sample
Simple random sample	Cluster sample	Quota sample
Systematic sample		

- (i) Three post codes in England are selected at random and all prices of houses for sale in those post codes are collected.
 - (ii) 50 boys and 50 girls are selected at random from the students in a college.
 - (iii) Every tenth person from those registered on a website is surveyed about his/her satisfaction with the website.
2. A school has 1205 students. Describe how to take a random sample of 50 students from the school.
3. A college wants to sample its students to find out how much they would be willing to pay for lunch.

Two possible sampling methods are suggested:

- Take a random sample of 100 students and email the survey to them.
- Put the survey on the college website and wait until 100 students have responded before closing the survey.

Which of these methods is likely to give the more representative sample? Explain your choice.

4. A researcher uses a snowball sampling method to find out when drug users first started taking drugs. The method is to first ask one drug user and then get each person surveyed to suggest other drug users to ask.
- (i) Is the sampling method biased? Give a reason for your answer.
 - (ii) Suggest a reason why this sampling method might be used.

Section 1: Collecting data

Exercise level 2

1. In the UK National Lottery, players choose six numbers from 1 to 59. There is a Lucky Dip option which allows the player to have six numbers chosen at random.
 - (i) The population being sampled from is the numbers 1 to 59. What is the population mean?
 - (ii) A particular Lucky Dip selection consists of the numbers
1, 5, 8, 11, 15, 18.
 - (A) What is the sample mean?
 - (B) Is this sample biased? Give a reason for your answer.
2. A researcher wants to find out what people who live in a particular town think about the local recycling facilities. Each of the following methods of taking a sample is biased. For each one, explain why it is biased.
 - (i) Ring people chosen at random from the telephone directory for the town.
 - (ii) Put an advert in the local paper asking for volunteers and choose at random from the volunteers.
 - (iii) Choose all people who come out of the local supermarket between 9am and 10am on a Monday.
3. The NHAMCS survey collects information from a large random sample of hospital records for visits to emergency departments in the United States. The survey shows that 15.9% of such visits are by patients aged 65 and over. Four possible conclusions based on this information are given below. For each conclusion, decide whether it is justified and give a reason for your answer.
 - (i) About 16% of the population of the United States is aged 65 and over.
 - (ii) About 16% of all visits to hospital emergency departments in the United States are made by people aged 65 and over.
 - (iii) About 16% of all visits to hospital emergency departments in the UK are made by people aged 65 and over.
 - (iv) About 8% of all visits to hospital emergency departments in the United States are made by females aged 65 and over.
4. An inspector is responsible for inspecting the care homes for the elderly in a particular area. There are 500 care homes in the area. The inspector can inspect one care home each day. She has the right to visit any care home in the area without giving notice of her visit. For each of the following methods of deciding which care homes to visit, give one advantage and one disadvantage of the method.
 - (i) List the care homes in alphabetical order and visit them in that order, one each day.
 - (ii) Each morning, choose a care home at random to visit.
 - (iii) If there have been complaints about any care home, visit it as soon as possible. On days when there are no care homes with complaints to visit, choose a care home at random from those which have not been inspected in the past year.

Section 1: Collecting data

Exercise level 3

1. The population of Tower Hamlets aged 16 and over is shown in the table below.

Age	Number
16 to 17	4,953
18 to 19	7,010
20 to 24	30,818
25 to 29	40,157
30 to 44	70,245
45 to 59	29,337
60 to 64	5,863
65 to 74	8,169
75 to 84	5,611
85 to 89	1,256
90 and over	534
Total	203,953

- (i) A researcher wants to take a sample of 200 people aged 16 and over from Tower Hamlets. Calculate how many should be taken from each of the 11 groups in the table to get a sample where the number in each group is proportional to the population.
- (ii) (A) Explain why the researcher might want to sample more people from each of the 85 to 89 and 90 and over groups than your figures in (i) suggest.
- (B) The researcher asks whether people are satisfied with recycling facilities in their area. She samples 5 people from each of the 85 to 89 and 90 and over groups. What should she multiply the number who say “yes” by in each of these groups to make the overall sample representative of the population?
2. A website which sells books allows customers to rate books from 1 star to 4 stars. 4 stars is the top rating. The ratings for the three books in a fiction series are shown below.

	Book 1	Book 2	Book 3
4 star	45	38	34
3 star	47	10	9
2 star	15	1	0
1 star	15	0	0

- (i) What type of sample of all readers of each book are those who rate the books online?
- (ii) Decide whether each of the following statements is true and justify your decisions.
- (A) The mean rating is increasing as the series goes on.

AQA Collecting and interpreting data

(B) This shows that the books are getting better.

3. A researcher wants to find out the average number of occupants in a car on a particular road. Two different sampling methods are used to estimate this number.

Method 1: There are an estimated 10 000 cars on the road in a week. 500 random numbers from 1 to 10 000 are chosen. A camera photographs the cars corresponding to these random numbers. The researcher counts the occupants in each car and the average is taken.

Method 2: The road leads to a large shopping centre. At 15 different times in the week, people ask all those going into the shopping centre to tell them how many occupants were in the car they came in. The average is taken.

- (i) Method 2 produces a larger average than method 1. Explain why.
- (ii) Which method is likely to give the best estimate of the average number of occupants in a car on the road? Justify your choice.

